

Parameter Value Selection: A Glimpse into Curve Fitting Bias

Parameter/Indicator values refers to the values that your extern variables take on. We are taking about the period values that your SMAs, ATR and other indicators take up [Eg. SMA(20), SMA(100)].

Importance of Parameter Value

When we create the first draft of our robot, we need to decide what initial parameters/indicator values to use. These values will affect the backtest results of our robot. For example, the choice between SMA(20) or SMA(100) will make a very big difference in the backtest results. At this point of time, choosing the exact optimal values is not important. However, we still need to choose values that are appropriate for our robots.

The screenshot shows a window titled 'Belinda_Complete' with a green header bar. Below the header are three tabs: 'Testing', 'Inputs', and 'Optimization'. The 'Optimization' tab is selected. It contains a table with the following columns: Variable, Value, Start, Step, and Stop. The table lists 15 variables with their respective values, start values, step values, and stop values. At the bottom right of the window are buttons for 'Load', 'Save', 'OK', 'Cancel', and 'Reset'.

Variable	Value	Start	Step	Stop
☐ MagicNumber	12345	12345	0	0
☐ SignalMail	false	false	true	
☐ Lots	1.0	1.0	0.0	0.0
☐ Slippage	3	3	0	0
☐ UseStopLoss	true	false	true	
☐ StopLoss	30	30	0	0
☐ UseTakeProfit	false	false	true	
☐ TakeProfit	0	0	0	0
☐ UseTrailingStop	true	false	true	
☐ TrailingStop	30	30	0	0
☐ sma_short	10	10	0	0
☐ sma_long	40	40	0	0
☐ atr_period	20	20	0	0
☐ atr_shift	11	11	0	0

How to select Initial Parameter Values

The initial values for our parameters has to be suitable for the inefficiency that we are trying to capture. Let's assume we are exploiting a trending inefficiency that is supposed to occur over many months, we are trading on 4hr charts and we use 2 SMAs in our trading to identify this inefficiency.

If we use SMA(10) and SMA(30), we are looking at a 40 hours (~ 1.5 days) and 120 hours (5 days) periods respectively. This is definitely not sufficient to identify a trending inefficiency of this magnitude/duration.

Our General Rule of Thumb

Alright so we get the idea that the initial parameter values are correlated with the duration of the inefficiency we are trying to capture. However, doing this requires a bit of trial and error and frankly, some experience.

Fortunately, we have a general rule of thumb:

Your lowest parameter value should be longer than the average holding period (average duration) of your trades.

There are some limitations:

- 1) This rule does not apply to all trading systems. Some trading systems are designed to hold trades of extremely long duration which are triggered by short term signals.
- 2) And yes, I know this is a chicken and egg issue as average holding period and parameter values are interrelated. For illustration:
 - a. In our first backtest, we have no prior knowledge of the average holding period. Hence, we need to make an intelligent guess.
 - b. Once we have the average holding period of our trades from our first backtest, we modify our parameter values.
 - c. We run our second backtest with the new values, we get a new average holding period. We modify our parameter values again.
 - d. We run our third backtest with the latest values, we get a new average holding period. We modify our parameter values again.
 - e. Repeat.

Thus when we choose parameter values, we need to go through this process of trial and error until we converge to a situation where either the rule of thumb holds, or a situation where you are simply comfortable with the parameter values.

Hence, this is an art and science (again).

Curve Fitting Bias

In this part, we informally give you a glimpse into Curve Fitting.

Definition of Curve Fitting:

The process of designing the trading system so closely to historical data that it becomes ineffective in the future.

Why? Because the future rarely reflects the past!

Because the future rarely reflects the past, we should not create a robot that is so closely pegged to the past. Doing so will cause the robot to perform badly when it encounters with new market environments in the future.

This is an important point to note. When we select the initial parameter values and do trial and error to find the best parameter values, we are Curve Fitting our robot (to an extent).

Ways to mitigate this:

- 1) Try to rely on experience to select parameter values that are suitable for the inefficiency at hand. Reduce trial and error.
- 2) During our optimisation process, we will conduct a procedure called Walk-Forward Optimisation. This procedure will minimise Curve Fitting. More on this in Chapter 34.

Is it possible to completely eliminate Curve Fitting?

We believe it is not¹. Two analogies:

- 1) The fact that we are designing a robot for an observed inefficiency (that occurred in the past) is Curve Fitting. For instance, we observe that the market trends upwards every last Friday of the month due to some seasonal effect. We then create a robot to exploit this inefficiency; this is Curve Fitting.
- 2) A manual trader trades based on his experience. This experience is based on past events/market behavior aka historical data. This is considered Curve Fitting as well.

Curve Fitting is inevitable. On the bright side, a small extent of Curve Fitting is not a bad thing. In fact, some Curve Fitting is required for us to start building robots, and this is acceptable.

¹ If your robot uses random entries and/or exits, is it considered Curve Fitting? Food for Thought